Computer Vision

Klaus Diepold
Lehrstuhl für Datenverarbeitung

23. Mai 2008

2 Bilderzeugung

2.1 Abbildung durch Linsen

Durch Betrachtung von ähnlichen Dreiecken in Bild 1 ergibt sich die Beziehung

$$\frac{1}{f} = \frac{1}{Z} + \frac{1}{z},$$

die die grundlegende Gleichung für die Abbildung mit dünnen Linsen darstellt. Der Punkt x ist das Bild des Punktes p und f beschreibt den Brennpunkt der Linse.

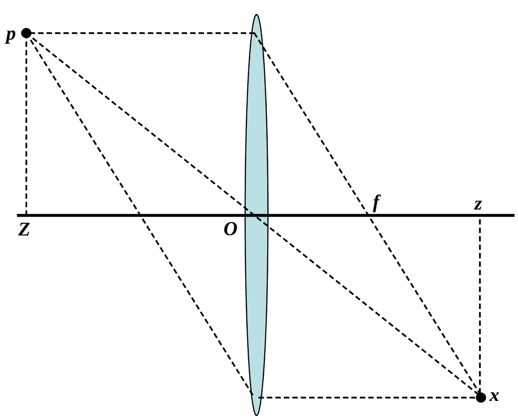

Bild 1: Abbildung mit dünnen Linsen.

2.2 Abbildung mittels Lochkamera

Bild 2 zeigt das Schema für die Abbildung des durch die Koordinaten $\boldsymbol{X} = \begin{bmatrix} X & Y & Z \end{bmatrix}^T$ gegebenen 3D-Raumpunktes P auf einen Punkt $\boldsymbol{x} = \begin{bmatrix} x & y \end{bmatrix}^T$ auf der Bildebene. Die Abbidung wird durch eine perspektifische Projektion

$$x = f \cdot \frac{X}{Z}, \quad y = f \cdot \frac{Y}{Z}$$

beschrieben. Der Parameter f bezeichnet dabei die Brennweite der Kamera. Wir verwenden oft eine vereinfachte Darstellung dieser Abbildung in Form einer Projektion π

$$\pi : \mathbb{R}^3 \to \mathbb{R}^2; \quad \boldsymbol{X} \mapsto \boldsymbol{x} = \pi(\boldsymbol{X}).$$

Die Projektion π wird als *ideale Lochkamera* bezeichnet. Der Bildpunktes $\boldsymbol{x}$ wird unter Verwendung

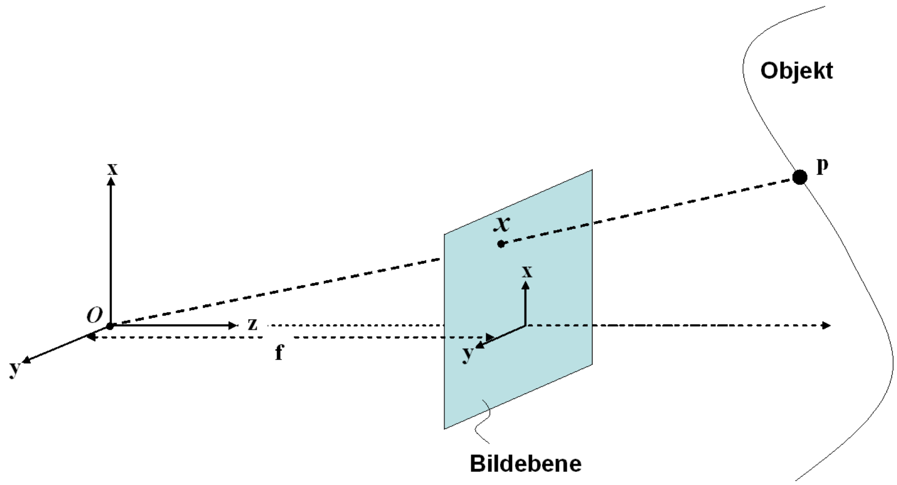

Bild 2: perspektifische (Loch-)Kamera.

homogener Koordinaten beschrieben als

$$\bar{\boldsymbol{x}} = \begin{bmatrix} fX/Z \\ fY/Z \\ 1 \end{bmatrix} \in \mathbb{R}^3.$$

2.3 Die ideale perspektifische Kamera

Wir betrachten einen generischen 3D-Punkt P mit den Koordinaten $\boldsymbol{X}_0 = \begin{bmatrix} X_0 & Y_0 & Z_0 \end{bmatrix}^T \in \mathbb{R}^3$, die bezogen auf ein Weltkoordinatensystem definiert sind. Der gleiche Punkt P wird im Koordinatensystem der Kamera dargestellt als $\boldsymbol{X} = \begin{bmatrix} X & Y & Z \end{bmatrix}^T$. Der Zusammenhang zwischen beiden Koordinatensystemen erfolgt mittels der Abbildung von $\boldsymbol{X}_0$ durch die Transformation $g(R, T)$ in der Form

$$\boldsymbol{X} = R\boldsymbol{X}_0 + T \quad \in \mathbb{R}^3.$$

Der Punkt P wird auf die Bildebene abgebildet (ideale Lochkamera)

$$\boldsymbol{x} = \begin{bmatrix} x \\ y \end{bmatrix} = \frac{f}{Z} \begin{bmatrix} X \\ Y \end{bmatrix}.$$

Unter Verwendung von homogenen Koordinaten wird diese Abbildung geschrieben als

$$Z \begin{bmatrix} x \\ y \\ 1 \end{bmatrix} = \left[\begin{array}{ccc|c} f & 0 & 0 & 0 \\ 0 & f & 0 & 0 \\ 0 & 0 & 1 & 0 \end{array} \right] \cdot \begin{bmatrix} X \\ Y \\ Z \\ 1 \end{bmatrix},$$

für den wir die verkürzte Schreibweise

$$Z\bar{\boldsymbol{x}} = \left[\begin{array}{ccc|c} f & 0 & 0 & 0 \\ 0 & f & 0 & 0 \\ 0 & 0 & 1 & 0 \end{array} \right] \cdot \bar{\boldsymbol{X}}$$

verwenden wollen. Hierbei sind $\bar{\boldsymbol{x}} = \begin{bmatrix} x & y & 1 \end{bmatrix}^T$ und $\bar{\boldsymbol{X}} = \left[\begin{array}{ccc|c} X & Y & Z & 1 \end{array} \right]^T$ die Darstellungen der Punkte $\boldsymbol{x}$ und $\boldsymbol{X}$ durch Verwendung von homogenenen Koordinaten.

2.4 Faktorisierung der Kamera-Matrix

Wir betrachten die Faktorisierung der Kamera-Matrix in der Form

$$\left[\begin{array}{ccc|c} f & 0 & 0 & 0 \\ 0 & f & 0 & 0 \\ 0 & 0 & 1 & 0 \end{array}\right] = \left[\begin{array}{ccc} f & 0 & 0 \\ 0 & f & 0 \\ 0 & 0 & 1 \end{array}\right] \cdot \left[\begin{array}{ccc|c} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{array}\right]$$

in die Faktoren

$$\boldsymbol{K}_f = \left[\begin{array}{ccc} f & 0 & 0 \\ 0 & f & 0 \\ 0 & 0 & 1 \end{array}\right] \in \mathbb{R}^{3\times 3}, \quad \Pi_0 = \left[\begin{array}{ccc|c} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{array}\right] \in \mathbb{R}^{3\times 4}.$$

Die Matrix Π_0 wird als kanonische Projektionsmatrix bezeichnet. Aus der Transformation zwischen Welt- und Kamerakoordinaten haben wir die Beziehung

$$\left[\begin{array}{c} X \\ Y \\ Z \\ 1 \end{array}\right] = \left[\begin{array}{cc} \boldsymbol{R} & \boldsymbol{T} \\ 0 & 1 \end{array}\right] \cdot \left[\begin{array}{c} X_0 \\ Y_0 \\ Z_0 \\ 1 \end{array}\right]$$

Die reelwertige 3×3-Matrix $\boldsymbol{R}$ beschreibt eine Rotation (Orientierung) der Kamera im 3D-Raum. Diese Rotation wird durch eine orthogonale Matrix beschrieben, die die folgenden Eigenschaften aufweist.

$$\boldsymbol{R}^T\boldsymbol{R} = \boldsymbol{R}\boldsymbol{R}^T = \mathbf{1}_3, \quad \text{und} \quad det\ (\boldsymbol{R}) = 1.$$

Allgemein hat eine orthogonale $n \times n$-Matrix $n(n-1)/2$ freie Paramter. Die vorliegende Matrix $\boldsymbol{R}$ hat somit 3 Freiheitsgrade, z.B. die drei Eulerwinkel. Die Menge aller 3×3-Rotationsmatrizen bilden eine Gruppe, die so genannte *Spezielle Orthogonale Gruppe*, die mit dem Symbol $SO(3)$ repräsentiert wird. Der reelwertige Vektor $\boldsymbol{T} \in \mathbb{R}^3$ beschreibt eine Translation der Kamera in den frei Raumrichtungen und weist ebenso 3 Freiheitsgrade auf. Die Größen $\boldsymbol{R}$ und $\boldsymbol{T}$ beschreiben zusammen die möglichen Bewegungen in 6 Freiheitsgraden für eine Kamera im 3D-Raum. Die Menge aller Transformationmatrizen der Form

$$\bar{\boldsymbol{g}} = \left[\begin{array}{cc} \boldsymbol{R} & \boldsymbol{T} \\ 0 & 1 \end{array}\right], \quad \boldsymbol{R} \in SO(3), \quad \boldsymbol{T} \in \mathbb{R}^3$$

bildet ebenfalls eine Gruppe, die so genannte *Spezielle Euklidische Gruppe*, die mit dem Symbol $SE(3)$ gekennzeichnet wird.

Wir fassen die Faktoren der Kamera-Matrix mit der Koordinatentransformation zusammen und erhalten einen Zusammenhang zwischen den Bildkoordinaten und den Weltkoordinaten des Punktes P in der Form

$$\lambda \cdot \left[\begin{array}{c} x \\ y \\ 1 \end{array}\right] = \left[\begin{array}{ccc} f & 0 & 0 \\ 0 & f & 0 \\ 0 & 0 & 1 \end{array}\right] \cdot \left[\begin{array}{ccc|c} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{array}\right] \cdot \left[\begin{array}{cc} \boldsymbol{R} & \boldsymbol{T} \\ 0 & 1 \end{array}\right] \left[\begin{array}{c} X_0 \\ Y_0 \\ Z_0 \\ 1 \end{array}\right].$$

Dieser Ausdruck lässt sich in kompakter Schreibweise schreiben

$$\lambda\bar{\boldsymbol{x}} = \boldsymbol{K}_f \cdot \Pi_0 \cdot \bar{\boldsymbol{X}} = \boldsymbol{K}_f \cdot \Pi_0 \cdot \bar{\boldsymbol{g}} \cdot \bar{\boldsymbol{X}}_0.$$

In den beiden vorangegangenen Gleichungen taucht der Parameter λ auf. Dieser Parameter bringt zum Ausdruck, dass der Vektor $\left[\begin{array}{ccc} x & y & 1 \end{array}\right]^T$ aufgrund seiner projektiven Natur (homogene Koordinaten, siehe nächsten Abschnitt) bis auf einen skalaren Faktor bestimmt ist, d.h. es gilt $\lambda\bar{\boldsymbol{x}} \sim \bar{\boldsymbol{x}}$.

Wenn die Brennweite f bekannt ist, und somit zu 1 normalisiert werden kann, dann können wir diesen Ausdruck noch vereinfachen zu

$$\lambda\bar{\boldsymbol{x}} = \Pi_0 \cdot \bar{\boldsymbol{X}} = \Pi_0 \cdot \bar{\boldsymbol{g}} \cdot \bar{\boldsymbol{X}}_0$$

2.5 Kamera mit intrinsischen Parametern

In diesem Abschnitt beschreiben wir den Übergang von der Beschreibung der Bildpunktkoordinaten in Millimetern

$$\boldsymbol{x} = \begin{bmatrix} x & y \end{bmatrix}^T$$

zur Beschreibung in Größeneinheiten von Pixeln gemessen wird, d.h.

$$\boldsymbol{x}_s = \begin{bmatrix} x_s & y_s \end{bmatrix}^T.$$

Dazu skalieren wird zunächste das Bild in x- und in y-Richtung

$$\boldsymbol{x}_s = \begin{bmatrix} x_s \\ y_s \end{bmatrix} = \begin{bmatrix} s_x & 0 \\ 0 & s_y \end{bmatrix} \cdot \begin{bmatrix} x \\ y \end{bmatrix}$$

damit die Bildkoordinaten mit der Größe des Bildsensors übereinstimen. Anschließend verschieben wir den Bildursprung vom Zentrum des Bildes (ideales Bild) auf das linken oberen Ecke des Bildes. Dies entspricht einem Übergang zu Sensorkoordinaten durch den Verschiebung der Pixelkoordinaten um den Offset-Vektor $\begin{bmatrix} o_x & o_y \end{bmatrix}$, der die Koordinaten des *Hauptpunktes* beinhaltet. Der Hauptpunkt ist der Punkt an dem die z-Aches (optische Achse) die Bildebene schneidet. Somit gilt

$$x' = x_s + o_x, \quad y' = y_s + o_y.$$

Als Ergebnis erhalten wir die Beschreibung von Punkten auf der Bildebene in Pixel-Koordinaten

$$\bar{\boldsymbol{x}}' = \begin{bmatrix} x' \\ y' \\ 1 \end{bmatrix}$$

statt der normalisierten Bildkoordinaten

$$\bar{\boldsymbol{x}} = \begin{bmatrix} x \\ y \\ 1 \end{bmatrix}$$

. Der Zusammenhang ist bildlich in Bild3 dargestellt.

Der gesamte Prozess kann durch Verwendung homogener Koordinaten kompakt zusammengefasst werden

$$\bar{\boldsymbol{x}}' = \begin{bmatrix} x' \\ y' \\ 1 \end{bmatrix} = \begin{bmatrix} s_x & 0 & o_x \\ 0 & s_y & o_y \\ 0 & 0 & 1 \end{bmatrix} \cdot \begin{bmatrix} x \\ y \\ 1 \end{bmatrix}.$$

Falls die Pixel nicht exakt rechtwinkelig sind, dann wird dies durch den Parameter s_θ beschrieben. Die Größe $s_\theta = \cot\theta$ beschreibt dabei den Einfluß des Winkels θ zwischen den Axen x_s und y_s. Die Berücksichtigung von nicht-rechtwinkeligen Pixeln in Form des Parameters s_θ führt auf die Matrixbeschreibung

$$\boldsymbol{K}_s = \begin{bmatrix} s_x & s_\theta & o_x \\ 0 & s_y & o_y \\ 0 & 0 & 1 \end{bmatrix} \in \mathbb{R}^{3\times 3},$$

die den Zusammenhang zwischen normalisierten Bildkoordinaten und realen Koordinaten auf dem Sensor. Insgesamt ergibt sich dadurch eine kompakte Beschreibung der Abbildung des 3D-Punktes (gegeben in Kamerakoordinaten) auf den Sensor (gegeben in Pixelkoordinaten) in Form der Kameraabbildung

$$\lambda \cdot \begin{bmatrix} x' \\ y' \\ 1 \end{bmatrix} = \begin{bmatrix} s_x & s_\theta & o_x \\ 0 & s_y & o_y \\ 0 & 0 & 1 \end{bmatrix} \cdot \begin{bmatrix} f & 0 & 0 \\ 0 & f & 0 \\ 0 & 0 & 1 \end{bmatrix} \cdot \left[\begin{array}{ccc|c} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{array}\right] \begin{bmatrix} X \\ Y \\ Z \\ 1 \end{bmatrix}.$$

Wir teilen die Kameraabbildung in zwei Teile auf, und zwar

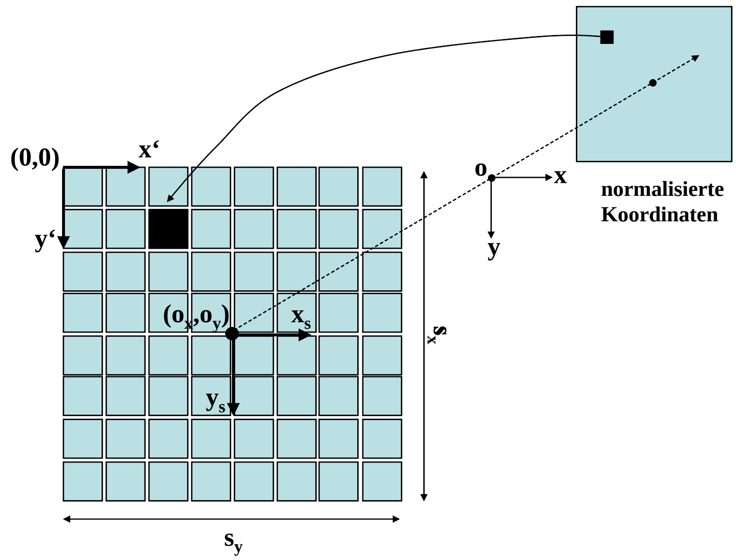

Bild 3: Intrinsische Parameter einer Kamera.

- in die perspektifische Projektion mit normalisiertem Koordinatensystem ($f = 1$), charakterisiert durch die kanonische Projektionsmatrix

$$\Pi_0 = \left[\begin{array}{ccc|c} 1 & & & 0 \\ & 1 & & 0 \\ & & 1 & 0 \end{array}\right]$$

- und die kameraspezifische Abbildung

$$\boldsymbol{K} = \boldsymbol{K}_s \cdot \boldsymbol{K}_f = \begin{bmatrix} s_x & s_\theta & o_x \\ 0 & s_y & o_y \\ 0 & 0 & 1 \end{bmatrix} \cdot \begin{bmatrix} f & 0 & 0 \\ 0 & f & 0 \\ 0 & 0 & 1 \end{bmatrix} = \begin{bmatrix} f \cdot s_x & f \cdot s_\theta & o_x \\ 0 & f \cdot s_y & o_y \\ 0 & 0 & 1 \end{bmatrix}.$$

Die obere Dreiecksmatrix $\boldsymbol{K}$ beinhaltet die *intrinsischen* Parameter der Kamera.

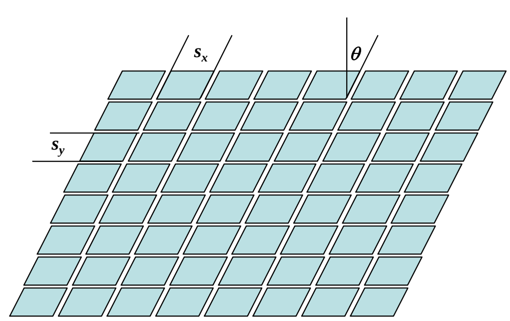

Bild 4: Pixel, deren Form nicht rechtwinkelig ist - Skew-Faktor.

Die Einträge der Matrix $\boldsymbol{K}$ haben folgende geometrische Bedeutung:

- o_x: x-Koordinate des Kamerahauptpunktes gemessen in Pixeln

- o_y: y-Koordinate des Kamerahauptpunktes gemessen in Pixeln
- $fs_x = \alpha_x$: Größe einer Längeneinheit, gemessen in horizontalen Pixeln
- $fs_y = \alpha_y$: Größe einer Längeneinheit, gemessen in vertikalen Pixeln
- α_x/α_y: Bildseitenverhältnis σ
- fs_θ: Skew, meist nahezu Null (siehe Bild 4)

Falls die intrinsischen Parameter der Kamera bekannt sind, d.h. wenn die Kamera intrinsisch kalibriert ist, dann können wir die normalisierten Koordinaten durch einfache Inversion berechnen

$$\lambda \boldsymbol{x} = \lambda \boldsymbol{K}^{-1}\boldsymbol{x'} = \Pi_0 \boldsymbol{X} = \left[\begin{array}{ccc|c} 1 & & & 0 \\ & 1 & & 0 \\ & & 1 & 0 \end{array}\right] \cdot \begin{bmatrix} X \\ Y \\ Z \\ 1 \end{bmatrix}.$$

Die Matrix $\boldsymbol{K}$ wird auch häufig als die Kalibrierungsmatrix bezeichnet. Der Prozess, die Kalibrierungsmatrix zu bestimmen wird das Kamera-Kalibrierung bezeichnet.

2.6 Elementare Geometrie

2.6.1 Punktraum und Vektorraum

Ein Punkt P im euklidischen Punktraum $\mathbb{E}^3$ ist beschrieben durch seine Ortskoordinaten $\boldsymbol{X}$

$$\boldsymbol{X} = \begin{bmatrix} X_1 \\ X_2 \\ X_3 \end{bmatrix} \in \mathbb{E}^3$$

Der Vektor $\boldsymbol{X}$ ist ein *gebundener* Vektor der die Lage von Punkten im $\mathbb{E}^3$ (Punktraum) beschreibt.

Definition 1 *Ein* freier *Vektor $\boldsymbol{v} \in \mathbb{V}^3$ ist bestimmt durch zwei Punkte $P, Q \in \mathbb{E}^3$ und zeigt von P nach Q. Der Punkt P hat die Koordinaten $\boldsymbol{X}$ und der Punkt Q die Koordinaten $\boldsymbol{Y}$. Dann ergibt sich $\boldsymbol{v}$ zu*

$$\boldsymbol{v} = \vec{PQ} = \begin{bmatrix} v_1 \\ v_2 \\ v_3 \end{bmatrix} = \boldsymbol{Y} - \boldsymbol{X} \in \mathbb{V}^3, \quad P, Q \in \mathbb{E}^3.$$

Freie Vektoren sind gegeben als Verbindungen zwischen Punkten im $\mathbb{E}^3$, d.h.

$$\boldsymbol{Y} - \boldsymbol{X} = \boldsymbol{Y'} - \boldsymbol{X'} \Rightarrow \quad \text{der selbe Vektor} \quad \boldsymbol{v}.$$

Die Menge der freien Vektoren bilden einen *lineraren Vektorraum* und ist eine Äquivalenzklasse von parallelen Pfeilen, bestimmt durch deren Länge und Richtung. Den Raum $\mathbb{V}^3$ bezeichnet man auch als den assoziierten Vektorraum. Die Summe aus einem Punkt P (gebundener Vektor) und einem freien Vektor $\boldsymbol{v}$ ergibt einen gebundenen Vektor (Punkt) $P - \boldsymbol{v} = Q$. Die Summe von Punkten (gebundenen Vektoren) ist nicht erklärt.

Im weiteren führen wir ein karthesisches Koordinatensystem ein , d.h. wir legen einen Ursprung $O \in \mathbb{E}^3$ fest (ausgezeichneter Punkt) sowie ein Orthonormalsystems $\{e_1, e_2, e_3\}$. Für die weiteren Betrachtungen können wir den Vektorraum $\mathbb{V}^3$ mit den Vektorraum $\mathbb{R}^3$ gleichsetzen.

Definition 2 *Im Vektorraum $\mathbb{V}^3$ ist ein* **Skalarprodukt** *definiert in der Form*

$$(\cdot, \cdot) : \mathbb{V}^3 \times \mathbb{V}^3 \to \mathbb{R}$$

$$(u, v) \mapsto u^T \cdot v = u_1 \cdot v_1 + u_2 \cdot v_2 + u_3 \cdot v_3, \quad \forall u, v \in \mathbb{V}^3$$

Definition 3 *Unter Verwendung des Skalarproduktes definieren wir eine* **Metrik** *in der Form*

$$\|u\|^2 = (u, u) = u^T \cdot u, \quad \forall u \in \mathbb{V}^3.$$

Definition 4 *Den* **Abstand** *zweier Punkte P und Q messen wir durch*

$$\|Y - X\|^2 = \|v\|^2.$$

Definition 5 *Für die Vektoren $\boldsymbol{u}, \boldsymbol{v} \in \mathbb{V}^3$ definieren wir das* **Kreuzprodukt**

$$\boldsymbol{u} \times \boldsymbol{v} = \begin{bmatrix} u_1 \\ u_2 \\ u_3 \end{bmatrix} \times \begin{bmatrix} v_1 \\ v_2 \\ v_3 \end{bmatrix} = \begin{bmatrix} u_2 \cdot v_3 - u_3 \cdot v_2 \\ u_3 \cdot v_1 - u_1 \cdot v_3 \\ u_1 \cdot v_2 - u_2 \cdot v_1 \end{bmatrix}.$$

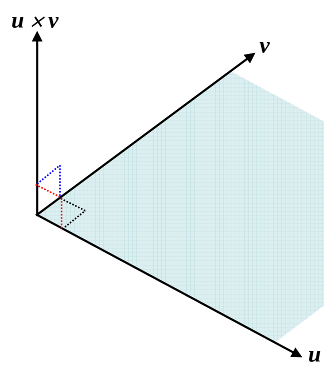

Bild 5: Kreuzprodukt im $\mathbb{V}^3$.

Das Kreuzprodukt weist eine Reihe wichtiger Eigenschaften auf, die im Folgenden zusammengestellt sind. Das Kreuzprodukt ist

- linear in beiden Argumenten: $\boldsymbol{u} \times (\alpha \boldsymbol{u} + \beta \boldsymbol{w}) = \alpha \boldsymbol{u} \times \boldsymbol{v} + \beta \boldsymbol{u} \times \boldsymbol{w}, \qquad \forall \alpha, \beta \in \mathbb{R}$
- orthogonal zu beiden Faktoren: $(\boldsymbol{u} \times \boldsymbol{v}, \boldsymbol{u}) = (\boldsymbol{u} \times \boldsymbol{v}, \boldsymbol{v}) = 0$ (siehe Bild 5)
- orientiert, d.h. es gilt $\boldsymbol{u} \times \boldsymbol{v} = -\boldsymbol{v} \times \boldsymbol{u}$. Die Vektoren $\boldsymbol{u}$, $\boldsymbol{v}$ und $\boldsymbol{u} \times \boldsymbol{v}$ bilden ein Rechtssystem (Rechte Hand Regel).

Wir können das Kreuzprodukt zwischen zwei Vektoren $\boldsymbol{u}$ und $\boldsymbol{v}$ als ein Matrix-Vektor-Produkt schreiben

$$\boldsymbol{u} \times \boldsymbol{v} = \hat{\boldsymbol{u}} \cdot \boldsymbol{v} = \begin{bmatrix} 0 & -u_3 & u_2 \\ u_3 & 0 & -u_1 \\ -u_2 & u_1 & 0 \end{bmatrix} \cdot \begin{bmatrix} v_1 \\ v_2 \\ v_3 \end{bmatrix} \in \mathbb{V}^3.$$

Im folgenden benutzen wir die Schreibweise

$$\hat{\boldsymbol{u}} = \begin{bmatrix} 0 & -u_3 & u_2 \\ u_3 & 0 & -u_1 \\ -u_2 & u_1 & 0 \end{bmatrix},$$

um die zum Kreuzprodukt gehörende Matrix $\hat{\boldsymbol{u}}$ zu repräsentieren. Die Matrix $\hat{\boldsymbol{u}}$ hat eine Reihe von Eigenschaften:

- Sie ist schief-symmterisch, d.h es gilt $\hat{\boldsymbol{u}}^T = -\hat{\boldsymbol{u}}$.
- Sie hat den Rang 2, d.h. es gilt $rank\,(\hat{\boldsymbol{u}}) = 2$.

- die Spalten der Matrix $\hat{\boldsymbol{u}}$ spannen den zu $\boldsymbol{u}$ orthogonalen 2-dimensionalen Unterraum auf.
- es gilt die Aussage: $\hat{\boldsymbol{u}}\boldsymbol{u} = 0$.
- Sie repräsentiert einen Isomorphismus zwischen der Menge aller schief-symmetrischen 3×3-Matrizen und dem Vektorraum $\mathbb{V}^3$. Die Menge aller schief-symmetrischen 3×3-Matrizen wird als *Lie-Algebra* bezeichnet und durch das Symbol $so(3)$ repräsentiert.
- Die Matrix erzeugt den Dach-Operator, der definiert ist durch die Vorschrift

$$\hat{(\cdot)} : \mathbb{V}^3 \to so(3); \quad \boldsymbol{u} \mapsto \hat{\boldsymbol{u}}.$$

- Der zugehörige inverse Dach-Operator ist somit gegeben durch die Vorschrift

$$(.)^\vee : so(3) \to \mathbb{V}^3; \quad \hat{\boldsymbol{u}}^\vee \mapsto \boldsymbol{u}.$$

2.6.2 Bild, Ur-Bild und Co-Bild

Wir betrachten die Abbildung der Geraden L im 3D-Raum auf die 2D-Bildebene (siehe Bild 6). Die Gerade L wird beschrieben durch einen Aufpunkt p_0 sowie einen Richtungsvektor $\boldsymbol{v}$. Der Aufpunkt p_0 und der Vektors $\boldsymbol{v}$ sind gegeben durch die (homogenen) Vektoren

$$\bar{\boldsymbol{X}}_0 = \begin{bmatrix} X_0 \\ Y_0 \\ Z_0 \\ 1 \end{bmatrix}, \quad \bar{\boldsymbol{V}} = \begin{bmatrix} V_1 \\ V_2 \\ V_3 \\ 0 \end{bmatrix}.$$

Damit kann jeder Punkt auf der Geraden L dargestellt werden durch

$$\bar{\boldsymbol{X}} = \bar{\boldsymbol{X}}_0 + \mu\bar{\boldsymbol{V}}, \quad \mu \in \mathbb{R}.$$

Dann ist das Bild der Geraden L gegeben durch die Menge aller Bildpunkte , die durch die homogenene Koordinaten

$$\bar{\boldsymbol{x}} \sim \Pi_0\bar{\boldsymbol{X}} = \Pi_0(\bar{\boldsymbol{X}}_0 + \mu\bar{\boldsymbol{V}}) = \Pi_0\bar{\boldsymbol{X}}_0 + \mu\Pi_0\bar{\boldsymbol{V}}.$$

beschrieben werden. Die Menge aller Punkte $\{\boldsymbol{x}\}$ spannen einen 2-dimensionalen Unterraum auf, sofern sie als Vektoren mit dem Fußpunkt O betrachtet werden. Der Schnitt dieses Unterraums (Ebene) mit der Bildebene ergbit eine Gerade als Schnittbild. Diese Schnittgerade ist das Bild der Geraden L. Die Frage ist dann, wie dieses Bild der Geraden am besten dargestellt werden kann.

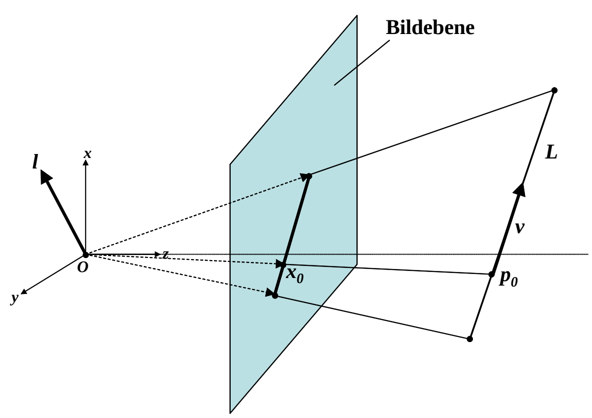

Bild 6: Perspektivisches Bild der Geraden L im 3D-Raum.

Definition 6 *Das* **Ur-Bild** *eines Punktes oder einer Geraden in der Bildebene ist definiert als die Menge der Punkte im 3-dimensionalen Raum, die auf diesen Punkt oder diese Gerade abgebildet werden.*

Definition 7 *Das* **Co-Bild** *eines Punktes oder einer Geraden ist definiert als der Unterraum des* $\mathbb{R}^3$*, der das (eindeutige) orthgonale Komplement des Ur-Bildes ist.*

Das Bild, das Ur-Bild sowie das Co-Bild sind alles äquivalente Repräsentationen in der Form, dass

$$Bild = Ur{-}Bild \cap Bildebene, \quad Ur{-}Bild = span\,(Bild), \quad Ur{-}Bild = Co{-}Bild^{\perp}, \quad Co{-}Bild = Ur{-}Bild^{\perp}$$

gilt. Das Ur-Bild der Geraden L ist ein 2-dimensionaler Unterraum. Deshalb ist das Co-Bild durch den Normalenvektor der Ebene repräsentiert. Die Notation dafür ist

$$\boldsymbol{l} = \begin{bmatrix} a \\ b \\ c \end{bmatrix} \in \mathbb{R}^3.$$

2.7 Projektive Geometrie und Transformationen in 2D

2.7.1 Darstelltung von Punkten und Geraden

Wir stellen eine Gerade als Nullstellenmenge der Gleichung

$$ax + by + c = 0$$

dar, oder auch in der Form

$$\boldsymbol{l}^T \bar{\boldsymbol{x}} = 0, \quad \boldsymbol{l} = \begin{bmatrix} a \\ b \\ c \end{bmatrix}, \quad \bar{\boldsymbol{x}} = \begin{bmatrix} x \\ y \\ 1 \end{bmatrix}.$$

Diese Darstellung ist nicht eindeutig weil für $k \neq 0$ die Gleichung

$$k\boldsymbol{l}^T \bar{\boldsymbol{x}} = kax + kby + kc = 0$$

die selbe Gerade beschreibt. Wir sprechen von einem *Projektiven Raum* $\mathbb{P}^2$, der äquivalent ist zum Raum $\mathbb{R}^3 - (0,0,0)^T$, d.h. ohne Urprung. Das bedeutet, daßein Projektiver Raum keinen Nullvektor enthält, oder anders ausgedrückt: alle Vektoren eines Projektiven Raumes haben mindestens eine Komponente verschieden von Null. Wir können den Punkt $\boldsymbol{x} \in \mathbb{V}^2$ durch Hinzfügen einer zusätzlichen Koordinate $z = 1$ durch *homogenen Koordinaten* darstellen, d.h. es gilt

$$\begin{bmatrix} x \\ y \end{bmatrix} \in \mathbb{V}^2 \quad \rightarrow \quad \begin{bmatrix} x \\ y \\ 1 \end{bmatrix} \in \mathbb{P}^2.$$

Der umgehkehrte Weg von homogonen Koordinaten zu inhomogenen Koordinaten erfolgt durch Division aller Koordinaten durch $z \neq 0$ und Streichen der dritten Koordinate, d.h.

$$\begin{bmatrix} x \\ y \\ z \end{bmatrix} \in \mathbb{P}^2 \rightarrow \begin{bmatrix} x/z \\ y/z \end{bmatrix} \in \mathbb{V}^2, \quad z \neq 0.$$

Durch Verwendung von homogenen Koordinaten ergeben sich folgende zwei Aussagen:

Satz 1 *Der Punkt* $\bar{\boldsymbol{x}} = \begin{bmatrix} x \\ y \\ 1 \end{bmatrix} \in \mathbb{P}^2$ *liegt auf der Geraden* $\boldsymbol{l} = \begin{bmatrix} a \\ b \\ c \end{bmatrix}$*, wenn* $\boldsymbol{l}^T \bar{\boldsymbol{x}} = 0$ *erfüllt ist.*

Satz 2 *Der Schnittpunkt zweier Geraden* $\boldsymbol{l}$ *und* $\boldsymbol{l}'$ *ist der Punkt* $\bar{\boldsymbol{x}} = \boldsymbol{l} \times \boldsymbol{l}' = \hat{\boldsymbol{l}} \cdot \boldsymbol{l}'$

Beispiel 1 *Schnittpunkt der beiden Geraden*

$$x = 1 \rightarrow \boldsymbol{l} = \begin{bmatrix} -1 \\ 0 \\ 1 \end{bmatrix}, \quad und \quad y = 1 \rightarrow \boldsymbol{l}' = \begin{bmatrix} 0 \\ -1 \\ 1 \end{bmatrix}$$

berechnet sich zu

$$\bar{\boldsymbol{x}} = \boldsymbol{l} \times \boldsymbol{l}' = \begin{bmatrix} -1 \\ 0 \\ 1 \end{bmatrix} \times \begin{bmatrix} 0 \\ -1 \\ 1 \end{bmatrix} = \begin{bmatrix} 0 & -1 & 0 \\ 1 & 0 & -1 \\ -0 & 1 & 0 \end{bmatrix} \cdot \begin{bmatrix} 0 \\ -1 \\ 1 \end{bmatrix} = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} \rightarrow \boldsymbol{x} = \begin{bmatrix} 1 \\ 1 \end{bmatrix}.$$

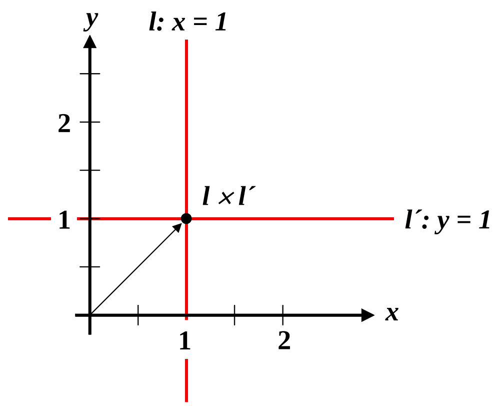

Bild 7: Schnitt zweier Geraden im $\mathbb{R}^2$.

Satz 3 *Die Verbindungsgerade* $\boldsymbol{l}$ *durch zwei gegebene Punkte* $\bar{\boldsymbol{x}}$ *und* $\bar{\boldsymbol{x}}'$ *ist gegeben durch* $\boldsymbol{l} = \bar{\boldsymbol{x}} \times \bar{\boldsymbol{x}}' = \hat{\bar{\boldsymbol{x}}} \cdot \bar{\boldsymbol{x}}'$

Beispiel 2 *Die beiden gegebenen Punkte*

$$\bar{\boldsymbol{x}} = \begin{bmatrix} 2 \\ 1 \\ 1 \end{bmatrix} \quad und \quad \bar{\boldsymbol{x}}' = \begin{bmatrix} 1 \\ 2 \\ 1 \end{bmatrix}$$

bestimmen die Gerade

$$\boldsymbol{l} = \bar{\boldsymbol{x}} \times \bar{\boldsymbol{x}}' = \begin{bmatrix} 2 \\ 1 \\ 1 \end{bmatrix} \times \begin{bmatrix} 1 \\ 2 \\ 1 \end{bmatrix} = \begin{bmatrix} 0 & -1 & 1 \\ 1 & 0 & -2 \\ -1 & 2 & 0 \end{bmatrix} \cdot \begin{bmatrix} 1 \\ 2 \\ 1 \end{bmatrix} = \begin{bmatrix} -1 \\ -1 \\ 3 \end{bmatrix}.$$

2.7.2 Ideale Punkte - Punkte im Unendlichen

Wir bestimmen den Schnitt der beiden Geraden, die durch die Gleichungen

$$ax + by + c = 0 \rightarrow \boldsymbol{l} = \begin{bmatrix} a & b & c \end{bmatrix}^T$$

und

$$ax + by + c' = 0 \rightarrow \boldsymbol{l}' = \begin{bmatrix} a & b & c' \end{bmatrix}^T$$

beschrieben sind. Der Schnittpunkt der beiden Geraden ergibt sich zu

$$\bar{\boldsymbol{x}} = \boldsymbol{l} \times \boldsymbol{l}' = \hat{\boldsymbol{l}} \cdot \boldsymbol{l}' = \begin{bmatrix} 0 & -c & b \\ c & 0 & -a \\ -b & a & 0 \end{bmatrix} \cdot \begin{bmatrix} a \\ b \\ c' \end{bmatrix} = \begin{bmatrix} bc' - bc \\ ac - ac' \\ 0 \end{bmatrix}.$$

Beim Übergang von homogenen Koordinaten zu euklidischen Koordinaten ergeben sich möglicherweise bei der Division mit 0 unendlich große Koordinatenwerte. Das bedeutet, dass der Scnittpunkt $\bar{\boldsymbol{x}}$ der beiden Geraden im Unendlichen liegt. Das wiederum bedeutet, dass die beiden Geraden parallel sind.

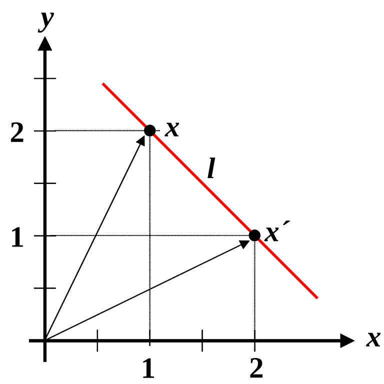

Bild 8: Geraden im $\mathbb{R}^2$ als Verbindung zweier Punkte.

Beispiel 3 *Der Schnittpunkt der beiden Geraden $x = 1$ und $x = 2$ berechnet sich mit*

$$\boldsymbol{l} = \begin{bmatrix} -1 \\ 0 \\ 1 \end{bmatrix}, \quad l' = \begin{bmatrix} -1 \\ 0 \\ 2 \end{bmatrix}$$

zu

$$\boldsymbol{l} \times \boldsymbol{l}' = \hat{\boldsymbol{l}} \cdot \boldsymbol{l}' = \begin{bmatrix} 0 & -1 & 0 \\ 1 & 0 & 1 \\ 0 & 1 & 0 \end{bmatrix} \cdot \begin{bmatrix} -1 \\ 0 \\ 2 \end{bmatrix} = \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix},$$

d.h. der Schnittpunkt liegt im Unendlichen in Richtung von y.

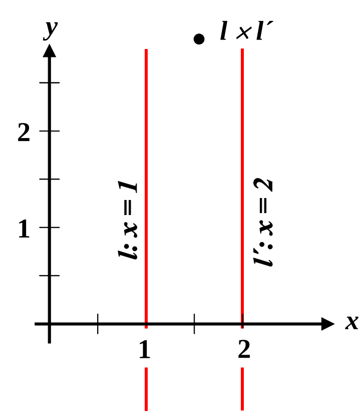

Bild 9: Schnittpunkt zweier paralleler Geraden im Unendlichen.

Die Punkte, die im Endlichen liegen und die durch homogene Koordinaten beschrieben bilden den Projektiven Raum, d.h. es gilt

$$\bar{\boldsymbol{x}} = \begin{bmatrix} x \\ y \\ z \end{bmatrix} \quad \text{mit} \quad z \neq 0 \quad \rightarrow \quad \text{Projektiver Raum} \quad \mathbb{P}^2.$$

Definition 8 *Punkte für die $z = 0$ erfüllt ist, d.h. für die $\bar{\boldsymbol{x}}_\infty = \begin{bmatrix} x & y & 0 \end{bmatrix}^T$ gilt, nennt man* ideale *Punkte.*

Die Gerade im Unendlichen ist gegeben durch $\boldsymbol{l}_\infty = \begin{bmatrix} 0 & 0 & 1 \end{bmatrix}^T$. Es gilt damit offensichtlich auch $\boldsymbol{l}_\infty^T \cdot \boldsymbol{x}_\infty = \begin{bmatrix} 0 & 0 & 1 \end{bmatrix}^T \cdot \begin{bmatrix} x \\ y \\ 0 \end{bmatrix} = 0$. Der Punkt im Unendlichen liegt auf der Geraden im Unendlichen. Die Gerade, die zwei Punkte im Unendlichen verbindet liegt ebenfalls im Unendlichen, z.B. betrachte man die beiden Punkte $\boldsymbol{x}_\infty$ und $\boldsymbol{x}'_\infty$ gegeben durch

$$\boldsymbol{x}_\infty = \begin{bmatrix} 2 \\ 1 \\ 0 \end{bmatrix}, \quad \boldsymbol{x}'_\infty = \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix}.$$

Die Verbindungsgerade zwischen diesen beiden Punkten berechnet sich zu

$$\boldsymbol{x}_\infty \times \boldsymbol{x}'_\infty = \hat{\boldsymbol{x}}_\infty \cdot \boldsymbol{x}'_\infty = \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} = \boldsymbol{l}_\infty$$

und liegt somit im Unendlichen.

2.7.3 Kegelschnitte

Im Rahmen der projektiven Geometrie sind alle Ellipsen unter projektiven Transformationen äquivalent, d.h. Ellipsen werden in Ellipsen abgebildet. Ellipsen werden uner Verwendung inhomogener Koordinaten beschrieben als Nullstellen eines Polynoms 2.Ordnung

$$ax^2 + bxy + cy^2 + dx + ey + f = 0.$$

unter Verwendung homogener Koordinaten $x \mapsto x_1/x_3$ und $y \mapsto x_2/x_3$ ergibt sich die Darstellung zu

$$ax_1^2 + bx_1x_2 + cx_2^2 + dx_1x_3 + ex_2x_3 + fx_3^2 = 0$$

oder in Matrixschreibweise

$$\bar{\boldsymbol{x}}^T \boldsymbol{C} \bar{\boldsymbol{x}} = 0$$

wobei die Matrix $\boldsymbol{C}$ gegeben ist durch

$$\boldsymbol{C} = \begin{bmatrix} a & b/2 & d/2 \\ b/2 & c & e/2 \\ d/2 & e/2 & f \end{bmatrix}.$$

Alle Punkte (x_i, y_i), die auf einer Ellipse liegen erfüllen die Gleichung

$$ax_i^2 + bx_iy_i + cy_i^2 + dx_i + ey_i + f = 0.$$

Dies kann auch geschrieben werden in der Form

$$\begin{bmatrix} x_i^2 & x_iy_i & y_i^2 & x_i & y_i & 1 \end{bmatrix} \cdot \boldsymbol{c} = 0,$$

wobei der Vektor $\boldsymbol{c} = \begin{pmatrix} a & b & c & d & e & f \end{pmatrix}^T$ die 6 Parameter der Ellipse beinhaltet. Diese Bedingung kann für 5 Punkte, die auf er Ellipse liegen, durch die entsprechenden Gleichungen zusammenfasst werden

$$\begin{bmatrix} x_1^2 & x_1y_1 & y_1^2 & x_1 & y_1 & 1 \\ x_2^2 & x_2y_2 & y_2^2 & x_2 & y_2 & 1 \\ x_3^2 & x_3y_3 & y_3^2 & x_3 & y_3 & 1 \\ x_4^2 & x_4y_4 & y_4^2 & x_4 & y_4 & 1 \\ x_5^2 & x_5y_5 & y_5^2 & x_5 & y_5 & 1 \end{bmatrix} \cdot \boldsymbol{c} = 0.$$

Die Ellipse wird somit beschieben als ein Vektor aus dem Kern dieser 5×6-Matrix.

Satz 4 *Die Gerade* $\boldsymbol{l}$*, die die Ellipse* $\boldsymbol{C}$ *im Punkt* $\boldsymbol{x}$ *berührt ist bestimmt durch die Bedingung* $\boldsymbol{l} = \boldsymbol{C}\boldsymbol{x}$.

Literatur

[1] Y. Ma, S. Soatto, J. Kosecka, S. Sastry. *An Invitation to 3-D Vision. From Images to Models.* Springer Verlag, 2006.

[2] R. Hartley, A. Zisserman. *Multiple View Geometry in Computer Vision.* Cambridge University Press, 2003.

[3] G. Strang. *Linear Algebra and its Applications.* Hartcourt Brace Jovanovich Publishers, San Diego, 1988.